Nama Kelompok:

1. Taufik Fathurahman (1301160790)
2. Haddad Alwi Yafie (1301162756)
3. Refka Muhammad Furqon (1301162748)

**Keterkaitan antara Tuberculosis dengan Angka Kematian Usia Dibawah Lima dan Usia Dewasa**

Dalam menentukan topik atau informasi yang akan digali dalam tugas besar penambangan data ini kami melakukan *data exploration* dengan melakukan visualisasi terhadap beberapa dataset yang telah diberikan. Dari dua belas dataset yang diberikan, kami menemukan tiga dataset yang memiliki korelasi yang kuat. Ketiga dataset tersebut adalah dataset *Tuberculosis incidence, Mortality rate, male adult,* dan *Mortality rate, under-five.*

|  |  |
| --- | --- |
| *Mortality rate, male adult* | *Mortality rate, under-five* |
|  |  |
| *Tuberculosis incidence* | |
|  | |

Table Correlation Matrix

Dari Table 1, didapatkan korelasi antar variable di satu dataset. Masing-masing variable memiliki keterkaitan yang kuat antar elemennya/variablenya. Lalu untuk melakukan identifikasi lebih lanjut, kami melakukan visualisasi untuk melihat korelasi antara tuberculosis dengan tingkat kematian anak dibawah lima tahun, dan usia dewasa pada Table 2.

|  |  |
| --- | --- |
| Tahun 2000 | Tahun 2005 |
|  |  |
| Tahun 2010 | Tahun 2011 |
|  |  |
| Tahun 2012 | Tahun 2013 |
|  |  |
| Tahun 2014 | Tahun 2015 |
|  |  |
| Tahun 2016 | |
|  | Keterangan Plot:  Merah: *Tuberculosis incidence*  Biru: *Mortality rate, under-five*  Hijau: *Mortality rate, male adult* |

Table Plot Korelasi

Dari table 2 kita bias melihat keterkaitan yang kuat, dimana ketika angka tuberculosis tinggi mengakibatkan angka kematian usia dibawah lima dan dewasa meningkat juga, begitu juga sebalaiknya bila angka tuberculosis menurun maka angka kematian usia dibawah lima dan dewasa menurun juga. Dari korelasi ini kami menyimpulkan untuk dapat melakukan pengambilan informasi yang akan di ambil dari keterkaitan ini untuk ditentukan algoritma yang cocok dalam menyelesaikan permasalahan pada tahapan selanjutnya.